

University of Groningen

The prediction of short- and long-term improvement in depressive patients

Bouhuys, Antoinette L.; Beersma, Domien G.M.; Hoofdakker, Rudi H. van den; Roossien, Albert

Published in:
Ethology and Sociobiology

DOI:
[10.1016/0162-3095\(87\)90024-0](https://doi.org/10.1016/0162-3095(87)90024-0)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1987

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bouhuys, A. L., Beersma, D. G. M., Hoofdakker, R. H. V. D., & Roossien, A. (1987). The prediction of short- and long-term improvement in depressive patients: Ethological methods of observing behavior versus clinical ratings. *Ethology and Sociobiology*, 8(8), 117S-130S. [https://doi.org/10.1016/0162-3095\(87\)90024-0](https://doi.org/10.1016/0162-3095(87)90024-0)

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The Prediction of Short- and Long-Term Improvement in Depressive Patients: Ethological Methods of Observing Behavior Versus Clinical Ratings

Antoinette L. Bouhuys, Domien G. M. Beersma,
Rudi H. van den Hoofdakker, and Albert Roossien
*Department of Biological Psychiatry, University Psychiatric Hospital,
Groningen, The Netherlands*

A considerable percentage of depressed patients do not respond to antidepressant treatment. Early indicators of prognosis clearly are needed. The aims of this study are to examine (1) whether the interpersonal behavior of patient and psychiatrist, as assessed by means of direct ethological observation of behavior during a clinical interview, might predict improvement, and (2) whether measures of psychomotor activation, as assessed by global clinical judgment, might predict improvement.

An analytical procedure was designed that allowed investigators to determine the organization of the observed interpersonal behavior. Speech-pause behavior was taken as the basic structure of the interaction, and the relationship of looking and of movements of hands and head to this structure was determined.

Baseline interviews conducted a few days after admission to the psychiatric hospital were studied on two groups of patients. In these groups, clinical improvement was predicted a posteriori over periods of 2 and 10 weeks of treatment, respectively.

Evidence is presented that improvement of depressed patients over both periods can be predicted on the basis of ethological methods of directly observing and analyzing behavior. In contrast, global clinical judgment of psychomotor activation, the method frequently used clinically, had no predictive potency.

KEY WORDS: Clinical interview; Clinical ratings; Depression; Dyadic interaction; Ethology; Gestures; Hand movements; Head movements; Looking; Observable behavior; Prediction; Speech-pause behavior.

Address reprint requests to: Dr. Antoinette L. Bouhuys, Department of Biological Psychiatry, University Psychiatric Hospital, Oostersingel 59, 9713 EZ Groningen, The Netherlands.

Ethology and Sociobiology 8: 117S-130S (1987)
© Elsevier Science Publishing Co., Inc., 1987
52 Vanderbilt Ave., New York, New York 10017

0162-3095/87/\$03.50

INTRODUCTION

Approximately 30% of depressed persons do not improve with pharmacological treatment (Klein 1980). Early prognostic indicators that would predict therapeutic effectiveness and/or spontaneous recovery are needed. Previous attempts to determine possible predictors of treatment outcome have focused mainly on clinical characteristics and psychobiological measures (Bielsky and Friedel 1976; Paykel 1979; Klein 1980).

On the diagnostic level, consensus exists that "endogenous" depressives have a better chance of improving on tricyclic antidepressants than have "neurotic" depressives. On the level of signs and symptoms constellations, strong evidence exists that within the endogenous type of depression the degree of psychomotor retardation (slowing down of thought processes and motoric behavior) is indicative of responsiveness to tricyclic antidepressants (Nelson et al. 1981). Data on the predictive properties of measures of anxiety are contradictory (Bielsky and Friedel 1976; Paykel 1979). As far as we know, only two studies have reported on prediction at the level of direct behavioral observation, description, and analysis. Ranelli and Miller (1981) suggest, for instance, that amounts of body touching (i.e., self-touching) during admission interviews are predictors of the response to antidepressive treatment with imipramine—smaller amounts of body touching predicting more improvement. In another study (Fossi et al. 1984), no behavioral indicators of treatment responsiveness to tricyclics were found, although 110 different behavioral elements were considered. In the current study, a new attempt was made to predict improvement in depression on the basis of ethological methods.

Direct observation of the behavior that the patient displays during a clinical interview, observation that includes listening to the patient's vocalizations, contributes to the psychiatrist's judgment about the patient's thoughts and feelings. Behavior, vocalization, thoughts (inferred from language), feelings (inferred from expressive behaviors), and longitudinal course are the components out of which diagnoses are made. The relatedness of these components prompted us to examine whether the *interaction* of patient and psychiatrist during such an interview might provide to an investigator observable predictors of improvement. Various considerations went into the determination of the choice of the elements to be studied and the analysis to be employed. For one thing, several studies have shown that the subjects' speech and looking behavior are the core behavioral elements for the judgment of interpersonal behavior, or social skill (Conger 1981; St. Lawrence 1982; Trower 1982). In addition, the relationship between verbal and visual behavior is important for the normal course of a conversation (Argyle et al. 1973; Kendon 1967).

Within the framework of studies on social skill, Trower (1982) suggested that in the investigation of the deficits in social-skill processes, more em-

phasis should be put on abnormal timing and monitoring of behaviors rather than on the search for unusual amounts of specific behaviors. A patient may exhibit behavior that differs in timing or monitoring from that of nonpatients while exhibiting no differences in the overall amount of the particular behavior. We designed an analytical procedure (Bouhuys and Alberts 1984) describing and quantifying the timing of various behaviors, such as looking, hand movements, and head movements, in relation to the speech-pause behavior of both participants of a conversation. As stated earlier, the clinical judgment of psychomotor retardation has been reported as predictive of treatment response (Nelson et al. 1981). Thus, objectively assessed behavioral components of this clinical concept of psychomotor retardation (or agitation) may provide predictors of clinical improvement, especially if aspects of timing are considered.

Apart from observed behavior, some psychometric test measures of psychomotor activation (i.e., retardation or agitation) were examined for their predictive potency as to therapeutic outcome.

METHODS

Subjects

The behavior of depressed inpatients with a psychiatrist was observed during baseline, medication-free clinical interviews. Two different groups of patients were studied and subjected to different designs (i.e., experiments). In experiment 1, clinical improvement was predicted a posteriori over a period of over 2 weeks ($n = 29$) (Bouhuys 1984), and in experiment 2, over 10 weeks ($n = 31$). The patients were included if they were diagnosed by an experienced psychiatrist as suffering from a major depressive disorder on the basis of *The Diagnostic and Statistical Manual III* (DSM-III 1980). The group of experiment 1 consisted of 23 women and 6 men [mean age 48 ± 14 (SD) years]. The group of experiment 2 consisted of 23 women and 8 men [mean age 51 ± 14 (SD) years]. During each patient's baseline interview, which was always conducted by the same, male psychiatrist, the degree of depression was assessed by at least two independent observers using the Hamilton Rating Scale for Depression (HRSD) (Hamilton 1967). The inter-rater reliability on the HRSD was 0.97 (Kendall concordance coefficient 0.97, $n = 28$) (Siegel 1956). A second individual interview with the patient was conducted after 2 weeks in experiment 1 and after 10 weeks in experiment 2. In the event that patients in experiment 2 were discharged earlier than 10 weeks, the interviews took place before discharge (range of experiment 2: 47–101 days). In these second interviews, the HRSD scores were measured again. Clinical change was assessed by subtracting the scores on the HRSD. On the basis of these differences, patients were classified as improved ($\text{HRSD} \geq 8$ units) or nonimproved ($\text{HRSD} < 8$ units) (Table 1).

Table 1. HRSD^a Scores for Improved and Nonimproved Depressed Patients After 2 or 10 Weeks of Treatment

	Improved (HRSD \geq 8)	Nonimproved (HRSD < 8)	Total
Experiment 1	(<i>n</i> = 15)	(<i>n</i> = 14)	(<i>n</i> = 29)
First interview (baseline)	30.7 \pm 7.0 ^b	29.8 \pm 6.4	30.3 \pm 6.5
Second interview (2 weeks)	17.3 \pm 7.6	26.5 \pm 6.4	
Experiment 2	(<i>n</i> = 18)	(<i>n</i> = 13)	(<i>n</i> = 31)
First interview (baseline)	29.4 \pm 5	27.7 \pm 5.8	28.7 \pm 5.3
Second interview (10 weeks)	5.9 \pm 3.4	23.8 \pm 5.8	

^a Hamilton Rating Scale for Depression (Hamilton, 1967).^b SD.

Patients were drug free at least 3 days before the start of the baseline interview. The 29 patients of experiment 1 were randomly assigned to three treatment groups. During a period of 2 weeks, they were treated with (1) clomipramine or (2) a combination of clomipramine and sleep deprivation or (3) a combination of placebo and sleep deprivation (Elsenga et al. 1983). After 2 weeks of treatment, no significant differences were found in the severity of depression between the treatment groups, although many patients in each group showed improvement. The 31 patients of experiment 2 received various treatments during their stay in the hospital, according to their individual clinical needs: tricyclics, neuroleptics, nonpharmacological therapies, and combinations of these. All therapeutic programs started between 2 and 3 weeks after the first drug-free interview. Item 8 of the HRSD, retardation, and item 9, agitation, are both rated on a 5-point scale. The interrater reliabilities on these items were 0.79 and 0.90, respectively (Kendall concordance coefficient). The mean baseline scores for retardation and agitation in experiment 1 were 1.2 ± 1.0 (SD) and 1.3 ± 1.0 (SD), respectively, and in experiment 2 were 1.1 ± 0.9 (SD) and 0.8 ± 0.7 (SD), respectively.

Observational Measures

The patients had given their written informed consent for video recordings of the interviews (split-screen technique). Two cameras were placed in such a way that both the patient and the psychiatrist could be viewed frontally. During the first 15 minutes (experiment 1) or 20 minutes (experiment 2), various behavioral categories of both persons were recorded continuously by means of an event-recording system. Registration of events was carried out in different runs of the videotape with preservation of temporal relationships. Changes in behavior were digitally coded to a precision of 0.2 seconds. The following behavioral categories were recorded:

1. *Sound Production* (speech) by both participants.
2. *Looking* by both persons, i.e., looking in the direction of the head of the partner.
3. *Hand Movements*: (a) *Body-touching hand movements* by the patient, i.e., movements with which the patient makes contact with any part of his or her body, such as touching or rubbing hair or face and manipulating the fingers. In experiment 2, a distinction has been made between *intensive* body touching [i.e., at least the forearm(s) and hand(s) are moving] and *light* body touching (i.e., only hands and/or fingers are moving). (b) *Object-touching hand movements* of the patient, i.e., manipulations of any object within the reach of the patient (e.g., plucking or rubbing a handbag, chair, or adornment). Also, for object-touching hand movements, a distinction has been made between an intensive and a light form, according to the same criteria. (c) *Gestures*, i.e., movements intimately linked to the rhythm and/or the content of speech.
4. *Head Movements*. (a) *Yes-nodding*. The presence was assessed on the basis of the content of the conversation and the frequency of the nodding. (b) *No-shaking*. (c) *Head movements*. All head movements not described under a and b.

The mean interrater reliability (kappa; Cohen 1968) for all scores was 0.89 (range 0.71–0.99).

Data Analysis

In the analysis of the organization of the behavior, two methods were employed.

Method 1. Behavior was studied in relation to speech-pause parameters. This method has been extensively presented elsewhere (Bouhuys and Alberts 1984). In short, to allow the exploration of the time relations between behavioral elements and speech-pause behavior of patient and psychiatrist, four types of pauses are distinguished, namely pauses within speech of either partner (speech pauses) and pauses during the switches of speech of the two partners in either direction (switching pauses) (Fig. 1a).

The analysis focuses on these four types of pauses and their preceding and following speech fragments. The sequence speech-pause-speech is called an *episode*. All speech-pause behavior displayed during an interview is subdivided into the four types of episodes (Fig. 1b: A, B, C, and D). The next step of the analysis is the description of the temporal distribution of other behavior over the four episodes. Figure 2 shows a model of the presentation of these distributions. The abscissa represents the time course of the four types of episodes. The hatched columns refer to speech of either person. Who is speaking is indicated above each of the columns. The distance between columns (within an episode) indicates the pause. This distance is pro-

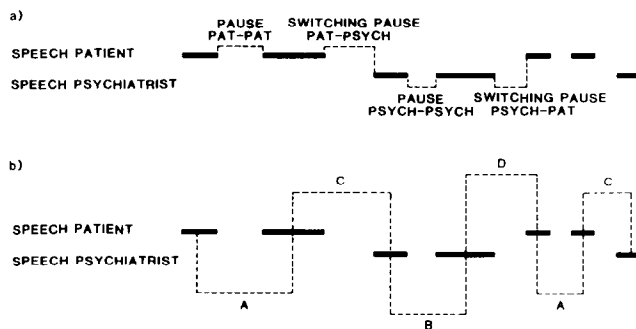


FIGURE 1. Illustration of methods (see text).

portional to the average duration of the specific pauses. The width of the columns gives the duration of the speech. The percentage of time that other behaviors (for instance, looking) occurred during the given episodes is represented on the ordinate. The fraction of episodes in which the behavior under consideration occurred at the specified time is indicated on the ordinate in Figures 3–6. In the current presentation a selection of episodes of specific durations was carried out. Only the so-called long episodes will be considered, having a speech duration of 0.4 to 10 seconds and a pause duration of 1 to 3 seconds. A normalization procedure was carried out to account for the variation in duration of speech and pause fragments.

The amount of looking by the psychiatrist was very high (96% percent of the observation time in experiment 1, and 87% of the observation time in experiment 2). Hence, looking behavior of the patients may be considered as patient and interviewer looking at each other. Method 1 has been carried out for looking and for the three types of hand movements.

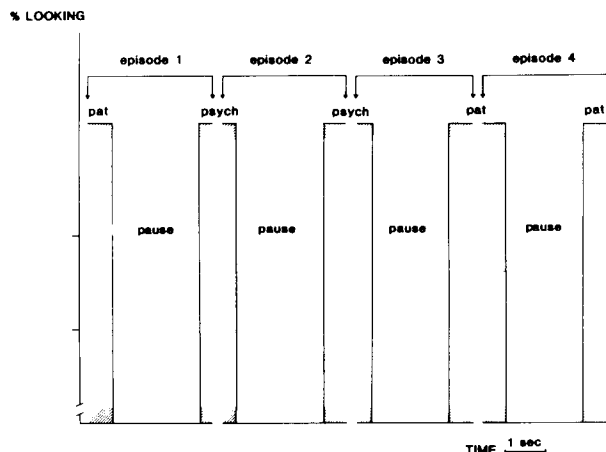


FIGURE 2. Illustration of methods (see text).

Method 2. The durations of the various behaviors relative to the observation time as well as their frequencies and mean durations were calculated.

Statistics

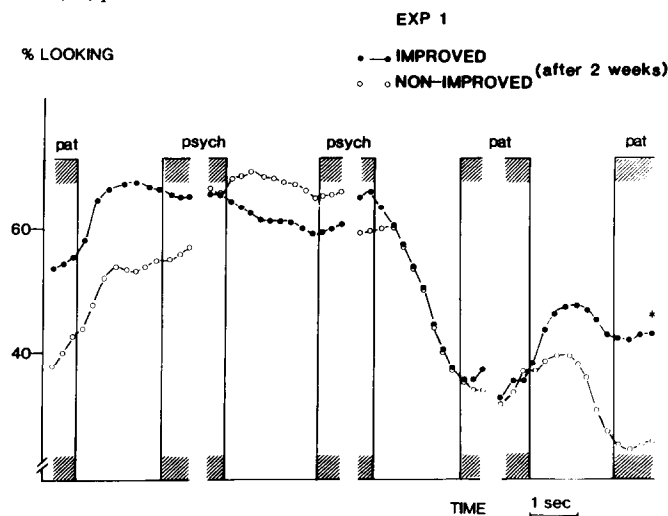
In experiment 1, Pearson correlations were calculated between looking and improvement. Comparisons between the groups comprising experiments 1 and 2, respectively, were made by means of the Mann-Whitney U test (Siegel 1956). For the method 1 analysis, group comparisons were made for the first, middle, and last moments of each episode. The significance level was set at $p < 0.05$ (two tailed).

RESULTS

Experiment 1 (Prediction Over 2 Weeks)

Looking. The distribution of looking during the baseline interview for improved and nonimproved patients is shown in Figure 3. The figure shows the characteristics of this behavior from the left to the right; at the end of the speaking turn of the patient (episode pat-psych), the patient frequently looks in the direction of the psychiatrist; during listening to the psychiatrist (episode psych-psych), the level of looking is relatively high; at the end of the psychiatrist's speaking turn, the patient usually looks away (episode psych-pat); during the patient's speaking turn, an increase of looking about halfway through the pause is found, which may be interpreted as a check

FIGURE 3. The amount of looking in relation to speech-pause-speech sequences (for the four types of episodes) during baseline interviews. Improved, $n = 15$; nonimproved, $n = 14$; *, $p < 0.05$.



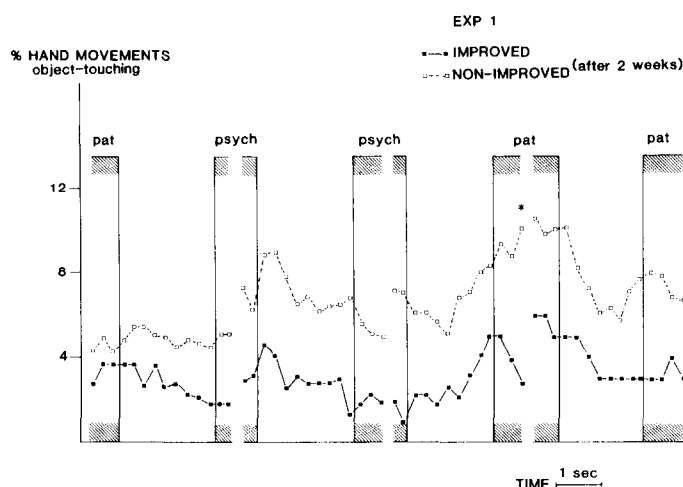


FIGURE 4. The amount of object-touching hand movements in relation to speech-pause-speech sequences (for the four types of episodes) during baseline interviews. Improved, $n = 15$; nonimproved, $n = 14$; *, $p < 0.02$.

on the listener's response. Moreover, with respect to the two groups, patients already behave differently during the baseline interview. At the end of the episode pat-pat, patients who will improve show higher amounts of looking than those who will not improve ($r = -0.47$; $n = 26$; $p < 0.05$). Other authors found that in normal subjects the ratio between the amount of looking while speaking and while listening was related to dominance, a high score indicating high levels of dominance (Dovidio and Ellyson, 1982). For the group of 29 patients, the ratio of the percent of looking during the last 0.2-second epoch of the episode psych-psych and the percent of looking during the last 0.2-second epoch of the episode pat-pat was calculated and correlated with the degree of subsequent improvement ($r = 0.55$; $n = 23$; $p < 0.01$). This relationship illustrates that improved patients show relatively high amounts of such dominance. (See Price and Sloman; Sloman and Price, both in this issue).

Object-touching hand movements. Figure 4 shows the distribution of object-touching hand movements (intensive and light together) for both groups of patients over the four types of episodes. Patients with relatively low amounts of object-touching hand movements at the start of their speaking turn, probably indicating low levels of arousal (Barroso et al. 1978; Mahl 1968), have more chance of improving than patients who show high amounts.

Body-touching hand movements and gestures. No differences were found between the groups with respect to the distribution of body-touching hand movements. Figure 5 shows the distribution of gestures. Again, the two groups behave differently during the interview; improved patients show

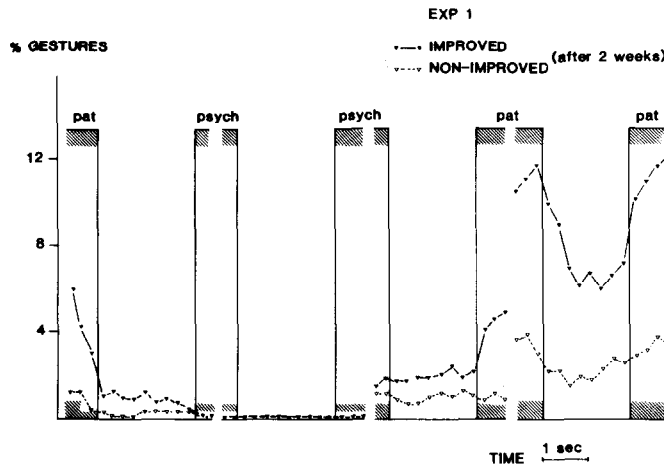


FIGURE 5. The amount of gestures in relation to speech-pause-speech sequences during baseline interviews. Improved, $n = 15$; nonimproved, $n = 14$; *, $p < 0.02$.

higher amounts of gestures than nonimproved patients. These higher amounts probably indicate more effort to communicate (Freedman 1972).

Other behaviors. The improved patients showed more yes-nodding than the non-improved ones, as assessed over the total observation time (method 2).

Experiment 2 (Prediction over 10 Weeks)

Looking. Figure 6 presents the distribution of looking for improved and nonimproved patients at the time of the baseline interview. No significant differences with respect to the amounts of looking could be established. However, visual inspection of these curves suggests a larger variation in amounts of looking during the episodes in the nonimprovers in comparison with improved patients. As a measure for this variation, the standard deviation of the amount of looking on the first and last moments of all four types of episodes has been calculated for each patient ($n = 8$ per patient). Figure 7 shows that this variation is significantly larger ($p < 0.02$) for non-improved patients.

Body-touching hand movements. With respect to body-touching hand movements, neither levels nor amount of variation over the episodes showed significant differences between the two groups. In contrast to experiment 1, in this experiment two levels of intensity of hand movements were differentiated. Figure 8 shows that the total duration of light and intensive body-touching hand movements are not different for both groups of patients. However, a considerable difference exists in the proportion of light and intensive body-touching hand movements between the two groups. The ratio of du-

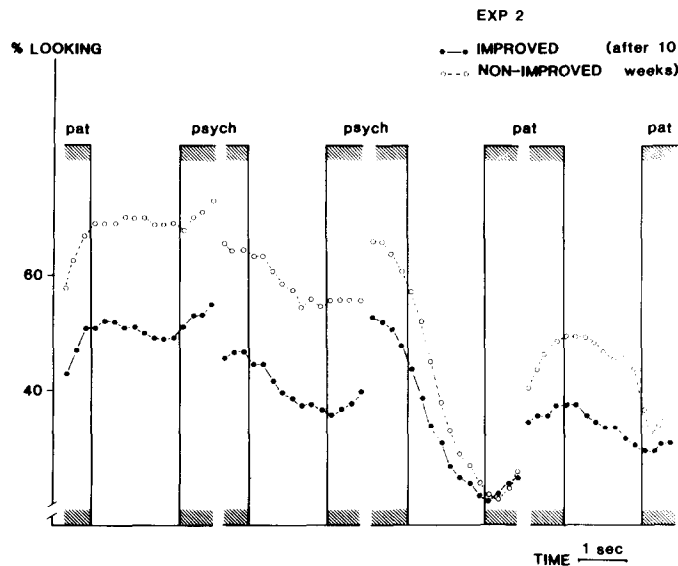


FIGURE 6. The amount of looking in relation to speech-pause-speech sequences during baseline interviews. Improved, $n = 18$; nonimproved, $n = 13$.

rations of light and intensive movements is significantly different ($p < 0.02$) between improved and nonimproved patients. Improved patients display relatively larger proportions of intensive body touching.

Object-touching hand movements and gestures. The distribution of object-touching hand movements and gestures is not different for the two groups. The total duration as well as the frequency of gestures is smaller in improved patients than in nonimproved patients (both $p < 0.05$).

VARIATION in intra-individual levels of looking

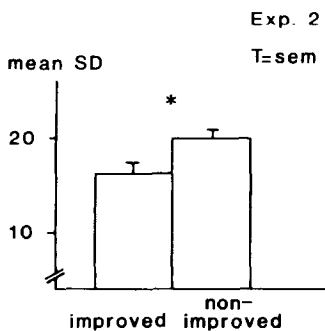


FIGURE 7. Variations in looking during the four types of episodes (SD of the amount of looking displayed on the first and last moments of all four types of episodes). Improved, $n = 18$; nonimproved, $n = 13$; *, $p < 0.02$.

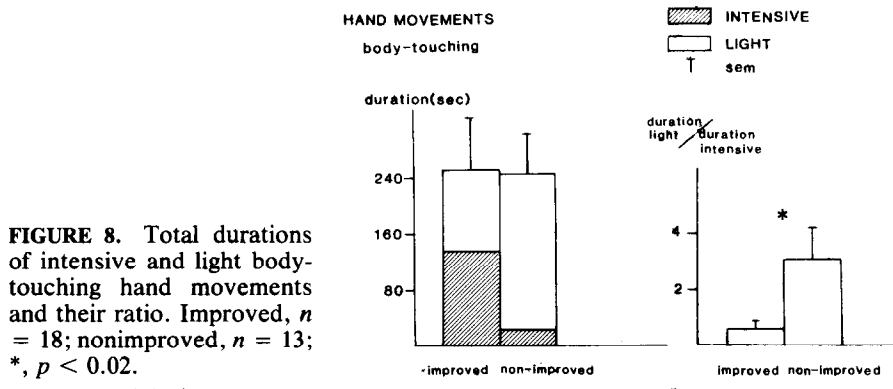


FIGURE 8. Total durations of intensive and light body-touching hand movements and their ratio. Improved, $n = 18$; nonimproved, $n = 13$; *, $p < 0.02$.

Head movements. The frequency of yes-nodding is significantly larger in nonimproved than in improved patients ($p < 0.05$), while the duration of the rest of their head movements is lower in comparison to improved patients ($p < 0.05$).

Retardation and agitation. In order to answer the question as to whether the state of psychomotor activation of patients could predict improvement, correlations between the degree of retardation and of agitation, rated on the HRSD, and the improvement of patients was calculated for both experiments. In both experiments, neither retardation nor agitation could predict improvement.

DISCUSSION

On the basis of directly observable behavior, improvement after 2 and after 10 weeks could be predicted. However, considerable differences exist in the types of behavioral variables having such predictive potency, probably indicating that different underlying processes are involved. After 2 weeks, those patients are improved who seem to agree with the psychiatrist (as reflected in more yes-nodding), who show more effort to communicate (i.e., make more gestures) and display more visual dominance (look more), and who are less aroused (as expressed by less object touching) during the interview. They probably are the patients who have high expectations of their admission, the ones who are more likely to react to improve irrespective of the type of treatment (experiment 1) and/or to show spontaneous remission. Would this relatively fast improvement after 2 weeks predict the state of depression after 10 weeks? Within a single group, no scores related to depression are available yet that were assessed after 2 and 10 weeks; nevertheless, the current data do not support this suggestion. On the contrary, most data assessed after 2 weeks are not in line with those assessed after 10 weeks

(looking, head movements). In fact, the two sets of data even point in the opposite direction (yes-nodding, gesturing).

Improved patients were considerably less depressive after 10 weeks (experiment 2) than after 2 weeks (experiment 1) (Table 1.) Therefore, prediction over a 10-week period is more relevant with respect to clinical practice and is considered now.

All behaviors observed result from an interaction between two persons. Various behaviors have regulative functions for or contribute to such interaction (like looking, yes-nodding, gesturing), whereas other behaviors are presumed to serve the regulation of the arousal state of the individual (e.g., body touching) and are probably less essential on the interactional level. The results are now discussed from this point of view.

Looking serves various functions in the interaction between people (Kendon 1967). For these functions to be realized in experiments 1 and 2, variations of levels over the various episodes are likely to be necessary (Bouhuys and Alberts 1984). In other words, variation in levels of looking is at least partially determined by the actions of the partner (the psychiatrist may for instance speak or keep silent). In the current study, this variation is small in improvers, which suggests that these patients are relatively independent of the actions of the psychiatrist. In other words, their behavior is relatively context independent. Also, other behaviors contributing to the interaction occur less during the interview—gesturing and yes-nodding, for instance. Both behaviors probably indicate less involvement in the interaction.

Hence, on the one hand, behaviors that contribute to the interaction occur *less* in patients who will improve. On the other hand, behaviors that probably serve the regulation of arousal states occur *more* in improvers; these individuals display high levels of intensive body touching and more head movements.

One may presume that this behavioral description of improvers reflects the global clinical concepts of psychomotor activation. No correlation data between retardation or agitation and directly observable behavior are available yet. However, it was shown that neither retardation nor agitation could predict improvement after 10 weeks.

On the global clinical-judgment level, these results seem not in agreement with other studies, in which the degree of retardation is found to be predictive for improvement on tricyclic medication (Paykel 1979; Nelson et al. 1981). Because of such matters as the large differences in procedures and the lack of significant information in the other studies, the discrepancy between these studies and ours cannot be explained.

No unequivocal data have been reported on the predictive ability of measures of anxiety (Bielsky and Friedel 1976). Supporting evidence that improvers are in states of higher arousal (as shown by more body touching and head movements) can be found in the study by Robin and Langley (1964). According to these authors, responders to imipramine had high anxiety and

emotional-lability ratings in comparison with nonresponders. Although in the current study more intensive body touching was indicative for improvement, in the Ranelli and Miller study the opposite was found, responders to imipramine showing less body touching than nonresponders (Ranelli and Miller 1981). On many points, the Ranelli and Miller study and our study are not comparable. For instance, Ranelli and Miller did not distinguish between two forms of body touching. In the current study, the intensive and light forms of body touching show opposite relationships with improvement. Hence, differences between the two studies in fraction of intensive and light body touching may account for differences in results. Also, in the Ranelli and Miller study, prediction over a period of 6 weeks was studied. In the current study, prediction over 2 and 10 weeks was studied, revealing different results between the two periods of time. Hence, differences in these periods also may account for the discrepancies found between the two studies.

In conclusion, evidence is presented that improvement of depressed patients could be predicted over periods of 2 and 10 weeks on the basis of ethological methods of directly observable behavior. Global clinical judgment of psychomotor activation, a method frequently used by clinicians, had no predictive potency. After 10 weeks of clinical treatment, consisting of such modes as psychotropic drugs, occupational therapy, and socio- and psychotherapy, those depressed patients improved who showed less interaction and a higher level of arousal during an interview shortly after admission.

REFERENCES

- Argyle, M., Alkema, F., Ingham, R., and McGallin, M. The different functions of gaze. *Semiotica* 7: 19–32, 1973.
- Barroso, F., Freedman, N., and Grand, S. Evocation of two types of hand movements in information processing. Human perception and performance. *Journal of Experimental Psychology* 4: 321–329, 1978.
- Bielsky, R.J., and Friedel, R.O. Predictions of tricyclic antidepressant response. A critical review. *Archives of General Psychiatry* 33: 1479–1489, 1976.
- Bouhuys, A.L. Observable behaviour as a predictor of improvement in endogenous depressive patients. In *Psychologische Aspekte depressiver Störungen*, M. Hautzinger and R. Straub (Eds.). Regensburg: Roderer Verlag, 1984, pp. 27–51.
- , Alberts, E. An analysis of the organization of looking and speech-pause behaviour of depressive patients. *Behaviour* 89: 270–298, 1984.
- Cohen, J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70: 213–220, 1968.
- Conger, J.C. Behavioral components of hetero-social skill. *Behaviour Therapy* 12: 41–55, 1981.
- Diagnostic and Statistical Manual III (DSM-III)*. Washington, D.C.: American Psychiatric Association Press, 1980.
- Dovidio, J.F., and Ellyson, S.R. Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly* 45: 106–113, 1982.

- Elsenga, S., and van den Hoofdakker, R.H. Clinical effects of sleep deprivation and clomipramine in endogenous depression. *Journal of Psychiatric Research* 17: 361–374, 1983.
- Fossi, L., Faravelli, C., and Paoli, M. The ethological approach to the assessment of depressive disorders. *Journal of Nervous and Mental Disease* 172: 332–341, 1984.
- Freedman, N. The analysis of movement behavior during the clinical interview. In *Studies in Dyadic Communication*, A.W. Siegman and B. Pope (Eds.). New York: Pergamon, 1972, pp. 153–175.
- Hamilton, M. Development of a rating scale for primary depressive illness. *British Journal of Clinical Psychology* 6: 278–296, 1967.
- Kendon, A. Some functions of gaze direction in social interaction. *Acta Psychologica* 26: 22–63, 1967.
- Klein, F. Review of the literature on mood-stabilizing drugs. In *Diagnosis and Drug Treatment of Psychiatric Disorders: Adults and Children*, 2nd ed., F. Klein, R. Gittelman, F. Quitkin, and A. Rifkin (Eds.). Baltimore: Williams & Wilkins, 1980, pp. 268–408.
- Mahl, J.F. Gestures and body movements in interviews. In *Research in Psychotherapy*, J.M. Shlien (Ed.). Washington: American Psychological Association, 1968, pp. 295–346.
- Nelson, J.C., Dennis, S., and Charney, S. The symptoms of major depressive illness. *American Journal of Psychiatry* 138: 1–13, 1981.
- Paykel, E.S. Predictions of treatment response. In *Psychopharmacology of Affective Disorders*, E.S. Paykel and A. Coppen (Eds.). New York: Oxford University Press, 1979, pp. 193–220.
- Ranelli, C.J., and Miller, R.E. Behavioral predictors of amitriptyline response in depression. *American Journal of Psychiatry* 138: 30–34, 1981.
- Robin, A.A., and Langley, G.E. A controlled trial of imipramine. *British Journal of Psychiatry* 110: 419–422, 1964.
- Siegel, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956, pp. 229–239.
- St. Lawrence, J.S. Validation of a component model of social skill with outpatient adults. *Journal of Behavioral Assessment* 4: 15–26, 1982.
- Trower, P. Situational analysis of the components and processes of behavior of socially skilled and unskilled patients. *Journal of Consulting and Clinical Psychology* 48: 327–339, 1982.